



IFWO

## RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/766,102

TIME: 14:43:58

Input Set : N:\Crf3\RULE60\10766102.raw.txt

Output Set: N:\CRF4\08302004\J766102.raw

```

1 <110> APPLICANT: Schnable, Patrick S.
2     Liu, Feng
3     Fu, Yan
4 <120> TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING MULTIPLE
5     START CODONS AND HISTIDINE TAGS
6 <130> FILE REFERENCE: 08411-027001
7 <140> CURRENT APPLICATION NUMBER: US/10/766,102
8 <141> CURRENT FILING DATE: 2004-01-27
9 <150> PRIOR APPLICATION NUMBER: US/09/897,776
10 <151> PRIOR FILING DATE: 2001-06-29
11 <150> PRIOR APPLICATION NUMBER: US 09/732,990
12 <151> PRIOR FILING DATE: 2000-12-08
13 <150> PRIOR APPLICATION NUMBER: US 60/169,725
14 <151> PRIOR FILING DATE: 1999-12-08
15 <160> NUMBER OF SEQ ID NOS: 37
16 <170> SOFTWARE: FastSEQ for Windows Version 4.0
18 <210> SEQ ID NO: 1
19 <211> LENGTH: 93
20 <212> TYPE: DNA
21 <213> ORGANISM: Artificial Sequence
22 <220> FEATURE:
23 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
--> 24 <221> NAME/KEY: CDS
25 <222> LOCATION: (1)...(84)
--> 26 <221> CDS
27 <222> LOCATION: (88)...(93)
--> 28 <400> 1
29     aag ctt cac cac cat cat cat cac gca tca cca cca cca cga atc      48
30     Lys Leu His His His His His His Ala Ser Pro Pro Pro Pro Arg Ile
31         1             5             10             15
32     atc atc acc atc acc tcg agc gtc aca cta gct gag taa gca tgc      93
33     Ile Ile Thr Ile Thr Ser Ser Val Thr Leu Ala Glu     Ala Cys
34         20             25             30
36 <210> SEQ ID NO: 2
37 <211> LENGTH: 66
38 <212> TYPE: DNA
39 <213> ORGANISM: Artificial Sequence
40 <220> FEATURE:
41 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
42 <400> SEQUENCE: 2
43     gtaccacca ccatcatcat cacgcatcac caccaccacc acgcatcatc atcaccatca      60
44     cctcga                                           66
46 <210> SEQ ID NO: 3

```

ENTERED

## RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/766,102

TIME: 14:43:58

Input Set : N:\Crf3\RULE60\10766102.raw.txt

Output Set: N:\CRF4\08302004\J766102.raw

```

47 <211> LENGTH: 14
48 <212> TYPE: DNA
49 <213> ORGANISM: Artificial Sequence
50 <220> FEATURE:
51 <223> OTHER INFORMATION: linker
52 <400> SEQUENCE: 3
53     ctgcagcggc cgcg                                     14
55 <210> SEQ ID NO: 4
56 <211> LENGTH: 22
57 <212> TYPE: DNA
58 <213> ORGANISM: Artificial Sequence
59 <220> FEATURE:
60 <223> OTHER INFORMATION: linker
61 <400> SEQUENCE: 4
62     ctaggcgccg gcgacgtctc ga                             22
64 <210> SEQ ID NO: 5
65 <211> LENGTH: 16
66 <212> TYPE: DNA
67 <213> ORGANISM: Artificial Sequence
68 <220> FEATURE:
69 <223> OTHER INFORMATION: linker
70 <400> SEQUENCE: 5
71     ctagctgcag atatca                                     16
73 <210> SEQ ID NO: 6
74 <211> LENGTH: 16
75 <212> TYPE: DNA
76 <213> ORGANISM: Artificial Sequence
77 <220> FEATURE:
78 <223> OTHER INFORMATION: linker
79 <400> SEQUENCE: 6
80     agcttgatat ctgcag                                     16
82 <210> SEQ ID NO: 7
83 <211> LENGTH: 25
84 <212> TYPE: DNA
85 <213> ORGANISM: Artificial Sequence
86 <220> FEATURE:
87 <223> OTHER INFORMATION: primer for PCR
88 <400> SEQUENCE: 7
89     ccatcgatcc gagatagggt tgagt                         25
91 <210> SEQ ID NO: 8
92 <211> LENGTH: 20
93 <212> TYPE: DNA
94 <213> ORGANISM: Artificial Sequence
95 <220> FEATURE:
96 <223> OTHER INFORMATION: primer for PCR
97 <400> SEQUENCE: 8
98     acgagctcag gcagagacga                               20
100 <210> SEQ ID NO: 9
101 <211> LENGTH: 20

```

## RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/766,102

TIME: 14:43:58

Input Set : N:\Crf3\RULE60\10766102.raw.txt

Output Set: N:\CRF4\08302004\J766102.raw

```

102 <212> TYPE: DNA
103 <213> ORGANISM: Artificial Sequence
104 <220> FEATURE:
105 <223> OTHER INFORMATION: primer for PCR
106 <400> SEQUENCE: 9
107      acgagctcgc agagacgacg
109 <210> SEQ ID NO: 10
110 <211> LENGTH: 26
111 <212> TYPE: DNA
112 <213> ORGANISM: Artificial Sequence
113 <220> FEATURE:
114 <223> OTHER INFORMATION: primer for PCR
115 <400> SEQUENCE: 10
116      cctcgagtca cacaggaaac agctaa
118 <210> SEQ ID NO: 11
119 <211> LENGTH: 24
120 <212> TYPE: DNA
121 <213> ORGANISM: Artificial Sequence
122 <220> FEATURE:
123 <223> OTHER INFORMATION: primer for PCR
124 <400> SEQUENCE: 11
125      ggctagcagc tggttcctgt gtga
127 <210> SEQ ID NO: 12
128 <211> LENGTH: 18
129 <212> TYPE: DNA
130 <213> ORGANISM: Artificial Sequence
131 <220> FEATURE:
132 <223> OTHER INFORMATION: primer for PCR
133 <400> SEQUENCE: 12
134      gtggagcatc tggtcgca
136 <210> SEQ ID NO: 13
137 <211> LENGTH: 37
138 <212> TYPE: DNA
139 <213> ORGANISM: Artificial Sequence
140 <220> FEATURE:
141 <223> OTHER INFORMATION: primer for PCR
142 <400> SEQUENCE: 13
143      gagatctgcc ataacatgtc atcatagctg tttcctg
145 <210> SEQ ID NO: 14
146 <211> LENGTH: 35
147 <212> TYPE: DNA
148 <213> ORGANISM: Artificial Sequence
149 <220> FEATURE:
150 <223> OTHER INFORMATION: linker
151 <400> SEQUENCE: 14
152      ctagccgaaa ttaatacgac tcactatagg gagac
154 <210> SEQ ID NO: 15
155 <211> LENGTH: 66
156 <212> TYPE: DNA

```

## RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/766,102

TIME: 14:43:58

Input Set : N:\Crf3\RULE60\10766102.raw.txt

Output Set: N:\CRF4\08302004\J766102.raw

```

157 <213> ORGANISM: Artificial Sequence
158 <220> FEATURE:
159 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
160 <400> SEQUENCE: 15
161     tatacatatg gcatggcatg gccactgcag gatccaccac catcatcatc acgcatcacc    60
162     accacc                                                                    66
164 <210> SEQ ID NO: 16
165 <211> LENGTH: 67
166 <212> TYPE: DNA
167 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
170 <400> SEQUENCE: 16
171     gacgtcgcat gcttactcag ctagtgtgat ggtgatgatg atggcctatg gtgggtggtgg    60
172     tgatgcg                                                                    67
174 <210> SEQ ID NO: 17
175 <211> LENGTH: 97
176 <212> TYPE: DNA
177 <213> ORGANISM: Artificial Sequence
178 <220> FEATURE:
179 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
180 <400> SEQUENCE: 17
181     taatacgact cactataggg agaccacaac ggtttccctc tagaaataat tttgtttaac    60
182     tttaagaagg agatatacat atggcatggc atggcca                                97
184 <210> SEQ ID NO: 18
185 <211> LENGTH: 13
186 <212> TYPE: DNA
187 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
190 <400> SEQUENCE: 18
191     atggcatggc atg                                                            13
193 <210> SEQ ID NO: 19
194 <211> LENGTH: 35
195 <212> TYPE: DNA
196 <213> ORGANISM: Artificial Sequence
197 <220> FEATURE:
198 <223> OTHER INFORMATION: linker
199 <400> SEQUENCE: 19
200     aattgtctcc ctatagtgag tcgtattaat ttcgg                                35
202 <210> SEQ ID NO: 20
203 <211> LENGTH: 28
204 <212> TYPE: PRT
205 <213> ORGANISM: Artificial Sequence
206 <220> FEATURE:
207 <223> OTHER INFORMATION: Synthetically generated peptide
208 <400> SEQUENCE: 20
209     Lys Leu His His His His His Ala Ser Pro Pro Pro Pro Arg Ile
210     1           5           10           15

```

## RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/766,102

TIME: 14:43:58

Input Set : N:\Crf3\RULE60\10766102.raw.txt

Output Set: N:\CRF4\08302004\J766102.raw

```

211      Ile Ile Thr Ile Thr Ser Ser Val Thr Leu Ala Glu
212                      20                      25
214 <210> SEQ ID NO: 21
215 <211> LENGTH: 93
216 <212> TYPE: DNA
217 <213> ORGANISM: Artificial Sequence
218 <220> FEATURE:
219 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
W--> 220 <221> NAME/KEY: CDS
      221 <222> LOCATION: (2)...(76)
W--> 222 <221> CDS
      223 <222> LOCATION: (80)...(91)
W--> 224 <400> 21
225      a agc ttc acc acc atc atc atc acg cat cac cac cac cac cac gca tca      49
226      Ser Phe Thr Thr Ile Ile Thr His His His His His Ala Ser
227      1                      5                      10                      15
228      tca tca cca tca cct cga gcg tca cac tag ctg agt aag cat      91
229      Ser Ser Pro Ser Pro Arg Ala Ser His      Leu Ser Lys His
230                      20                      25
231      gc      93
233 <210> SEQ ID NO: 22
234 <211> LENGTH: 25
235 <212> TYPE: PRT
236 <213> ORGANISM: Artificial Sequence
237 <220> FEATURE:
238 <223> OTHER INFORMATION: Synthetically generated peptide
239 <400> SEQUENCE: 22
240      Ser Phe Thr Thr Ile Ile Thr His His His His His His Ala Ser
241      1                      5                      10                      15
242      Ser Ser Pro Ser Pro Arg Ala Ser His
243                      20                      25
245 <210> SEQ ID NO: 23
246 <211> LENGTH: 4
247 <212> TYPE: PRT
248 <213> ORGANISM: Artificial Sequence
249 <220> FEATURE:
250 <223> OTHER INFORMATION: Synthetically generated peptide
251 <400> SEQUENCE: 23
252      Leu Ser Lys His
253      1
255 <210> SEQ ID NO: 24
256 <211> LENGTH: 93
257 <212> TYPE: DNA
258 <213> ORGANISM: Artificial Sequence
259 <220> FEATURE:
260 <223> OTHER INFORMATION: Synthetically generated oligonucleotide
W--> 261 <221> NAME/KEY: CDS
      262 <222> LOCATION: (3)...(80)
W--> 263 <221> CDS

```

## VERIFICATION SUMMARY

DATE: 08/30/2004

PATENT APPLICATION: US/10/766,102

TIME: 14:43:59

Input Set : N:\Crf3\RULE60\10766102.raw.txt

Output Set: N:\CRF4\08302004\J766102.raw

L:24 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:26 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:1  
L:28 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:1  
L:220 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:222 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:21  
L:224 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:21  
L:261 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:263 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:24  
L:265 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:24  
L:302 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:304 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:27  
L:306 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:27  
L:346 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:348 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:30  
L:350 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:30  
L:352 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:30  
L:400 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:402 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:34  
L:404 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:34